

Calculating a maximizer for quantum mutual information

Tony Dorlas^{1,*} and Ciara Morgan^{1,†}

¹*School of Theoretical Physics
Dublin Institute for Advanced Studies
10 Burlington Road
Dublin 4, Ireland.*

(Dated: January 31, 2008)

We obtain a maximizer for the quantum mutual information for classical information sent over the quantum qubit amplitude damping channel. This is achieved by limiting the ensemble of input states to antipodal states, in the calculation of the product-state capacity for the channel, the resulting maximizing ensemble consisting of just two non-orthogonal states. We also consider the product-state capacity of a convex combination of two memoryless channels and demonstrate in particular that it is in general not given by the minimum of the capacities of the respective memoryless channels.

Keywords: product-state capacity; maximizing ensemble; memory channel.

I. INTRODUCTION

In this paper we obtain the product-state capacity of the amplitude damping channel. It is determined by a transcendental equation in a single real variable, which is easily solved numerically. We also consider a convex combination of two memoryless channels and show in particular that the product-state capacity of a convex combination of a depolarizing and an amplitude damping channel, which was shown in [1] to be given by the supremum of the minimum of the corresponding Holevo quantities, is not equal to the minimum of their product-state capacities.

A. Memoryless channels and the HSW theorem

The transmission of classical information over a quantum channel is achieved by encoding the information as quantum states. A memoryless channel is given by a completely positive trace-preserving map $\Phi : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{K})$, where $\mathcal{S}(\mathcal{H})$ and $\mathcal{S}(\mathcal{K})$ denote the states on the input and output Hilbert spaces \mathcal{H} and \mathcal{K} respectively. In the case of product-state inputs, the HSW theorem, proved independently by Holevo [2] and by Schumacher and Westmoreland [3], states that the *product-state capacity* for classical information sent through a memoryless quantum channel is given by

$$\chi^*(\Phi) = \max_{\{p_j, \rho_j\}} \chi(\Phi)(\{p_j, \rho_j\}), \quad (1)$$

where the Holevo- χ -quantity is defined by

$$\chi(\Phi)(\{p_j, \rho_j\}) := S\left(\sum_j p_j \Phi(\rho_j)\right) - \sum_j p_j S(\Phi(\rho_j)), \quad (2)$$

and where S is the von Neumann entropy, $S(\rho) = -\text{tr}(\rho \log \rho)$. The maximum is taken over all ensembles of input states ρ_j with probabilities p_j . The capacity for channels with entangled input states has been studied [4], and it has been shown that for certain channels the use of entangled states can enhance the inference of the output state and increase the capacity (e.g. [5]). We concentrate here on the *product-state* capacity for noisy quantum channels.

Note that, by concavity of the entropy, the maximum in Equation (1) is always attained for an ensemble of *pure states* ρ_j . Moreover, it follows from Carathéodory's theorem (see [6–8]), that the ensemble can always be assumed to contain no more than d^2 pure states, where $d = \dim(\mathcal{H})$.

*Electronic address: dorlas@stp.dias.ie

†Electronic address: cqtciaara@nus.edu.sg

In Section II we show that, in the case of the amplitude damping channel, the maximum is in fact obtained for an ensemble of two pure states [9]. Moreover, these states are in general not orthogonal as in the channel considered by Fuchs [11]. Figure 1 demonstrates the action of the amplitude damping channel with error parameter $\gamma = \frac{1}{2}$ with the optimal input-states represented in blue and the corresponding output states in red.

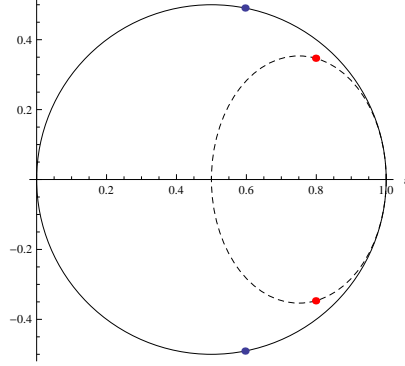


FIG. 1: Optimal input-states (blue) to the amplitude-damping channel with $\gamma = 0.5$ and the resulting output states from the channel (red).

B. Convex combination of memoryless channels

In [1] the product-state capacity of a convex combination of memoryless channels was determined. Given a finite collection of memoryless channels Φ_1, \dots, Φ_M with common input Hilbert space \mathcal{H} and output Hilbert space \mathcal{K} , a convex combination of these channels is defined by the map

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^M \gamma_i \Phi_i^{\otimes n}(\rho^{(n)}), \quad (3)$$

where γ_i , ($i = 1, \dots, M$) is a probability distribution over the channels Φ_1, \dots, Φ_M . Thus, a given input state $\rho^{(n)} \in \mathcal{S}(\mathcal{H}^{\otimes n})$ is sent down one of the memoryless channels with probability γ_i . This introduces long-term memory, and as a result the capacity of the channel $\Phi^{(n)}$ is no longer given by the maximum of the Holevo quantity. Instead, it was proved in [1] that it is given by

$$C_p(\Phi^{(n)}) = \sup_{\{p_j, \rho_j\}} \left[\bigwedge_{i=1}^M \chi_i(\{p_j, \rho_j\}) \right], \quad (4)$$

where $\chi_i = \chi(\Phi_i)$ is the Holevo quantity for the i -th channel Φ_i .

II. THE AMPLITUDE-DAMPING CHANNEL AND THE HOLEVO- χ -QUANTITY.

The qubit amplitude-damping channel models the loss of energy in a qubit quantum system and is described, with error parameter $0 \leq \gamma \leq 1$, by the following operation elements

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}. \quad (5)$$

Using the operation elements above, the qubit amplitude-damping channel can be expressed as follows $\Phi_{amp}(\rho) = E_0 \rho E_0^* + E_1 \rho E_1^*$. Note that since $E_0^* E_0 + E_1^* E_1 = I$, the operator Φ_{amp} is a CPT map and therefore a legitimate quantum channel.

Acting on the general qubit state $\rho = \begin{pmatrix} a & b \\ \bar{b} & 1-a \end{pmatrix}$, the amplitude-damping channel Φ_{amp} is given by

$$\Phi_{amp}(\rho) = \begin{pmatrix} a + (1-a)\gamma & b\sqrt{1-\gamma} \\ \bar{b}\sqrt{1-\gamma} & (1-a)(1-\gamma) \end{pmatrix}. \quad (6)$$

The eigenvalues of $\Phi_{amp}(\rho)$ are easily found to be

$$\lambda_{amp\pm} = \frac{1}{2} \left(1 \pm \sqrt{(1 + 2a(\gamma - 1) - 2\gamma)^2 - 4|b|^2(\gamma - 1)} \right). \quad (7)$$

To maximize the Holevo quantity, given by Equation (2), for this channel we show that the first term is increased, while keeping the second term fixed, if each pure state ρ_j is replaced by itself and its mirror image in the real b -axis, i.e. if we replace $\rho_j = \begin{pmatrix} a_j & b_j \\ b_j^* & 1 - a_j \end{pmatrix}$ associated with probability p_j , with the states $\rho_j = \begin{pmatrix} a_j & b_j \\ b_j^* & 1 - a_j \end{pmatrix}$ and $\rho'_j = \begin{pmatrix} a_j & -b_j \\ -b_j^* & 1 - a_j \end{pmatrix}$, both with probabilities $p_j/2$.

In general, the states ρ_j must lie inside the Poincaré sphere $(a - \frac{1}{2})^2 + |b|^2 \leq \frac{1}{4}$ and so the pure states will lie on the boundary $|b|^2 = a(1 - a)$.

We first show that the second term in Equation (2) remains unchanged when the states are replaced in the way described above. Indeed, since the eigenvalues (7) depend only on $|b|$, we have $S(\Phi(\rho_j)) = S(\Phi(\rho'_j))$ and therefore the first term is unchanged. Secondly, by concavity and the fact that $S(\sum_j p_j \Phi(\rho'_j)) = S(\sum_j p_j \Phi(\rho_j))$, we get,

$$S\left(\sum_j \frac{p_j}{2} \Phi(\rho_j + \rho'_j)\right) \geq S\left(\Phi\left(\sum_j p_j \rho_j\right)\right). \quad (8)$$

We can conclude that the first term in Equation (2) is increased with the second term fixed if each state ρ_j is replaced by itself together with its mirror image.

A. Convexity of the output entropy

We concentrate here on proving that, in the case of the amplitude-damping channel, the second term in the equation for the Holevo- χ -quantity is convex as a function of the parameters a_j when ρ_j is taken to be a pure state, i.e. $b_j = \sqrt{a_j(1 - a_j)}$. (Note that $S(a)$ only depends on $|b|$.) Thus $S(\Phi(\rho_j))$ is a function of one variable only, i.e. $S(a_j) = S(\Phi_{amp}(\rho_{a_j}))$, with $\rho_a = \begin{pmatrix} a & \sqrt{a(1 - a)} \\ \sqrt{a(1 - a)} & 1 - a \end{pmatrix}$ and hence

$$\sigma(a) = \Phi_{amp}(\rho_a) = \begin{pmatrix} \frac{a + (1 - a)\gamma}{\sqrt{a(1 - a)}\sqrt{1 - \gamma}} & \frac{\sqrt{a(1 - a)}\sqrt{1 - \gamma}}{(1 - a)(1 - \gamma)} \end{pmatrix}. \quad (9)$$

The eigenvalues of (9) are given by $\lambda_{amp\pm} = \frac{1}{2}(1 \pm x)$, where $x = \sqrt{1 - 4\gamma(1 - \gamma)(1 - a)^2}$, and thus $S(a) = H(\frac{1+x}{2})$, where $H(p) = -p \log p - (1 - p) \log(1 - p)$ is the binary entropy. It is now easy to see that $S''(a) \geq 0$ and hence that $S(a)$ is convex. Writing $\bar{\rho}_a = \sum_j p_j \rho_{a_j}$ with $\bar{a} = \sum_j p_j a_j$ and $\chi_{AD}(\{p_j, \rho_j\}) = \chi(\Phi_{amp})(\{p_j, \rho_j\})$ we have

$$\chi_{AD}(\{p_j, \rho_j\}) = S(\Phi_{amp}(\bar{\rho}_a)) - \sum_j p_j S(a_j) \leq S(\Phi_{amp}(\bar{\rho}_a)) - S(\bar{a}). \quad (10)$$

The capacity is therefore given by

$$\chi(\Phi_{amp}) = \max_{a \in [0, 1]} \left[S\left(\frac{1}{2}(\sigma(a) + \sigma'(a))\right) - S(\sigma(a)) \right]. \quad (11)$$

The maximizing value of a is given by the transcendental equation $\chi'_{AD}(a) = 0$ and can only be computed numerically.

It turns out that $a_{\max} \geq \frac{1}{2}$ for all γ . This is in fact easily proved: The determining equation is

$$\chi'_{AD}(a) \ln 2 = -(1 - \gamma) \ln \frac{a + \gamma(1 - a)}{(1 - \gamma)(1 - a)} + \frac{2\gamma(1 - \gamma)(1 - a)}{x} \ln \frac{1 + x}{1 - x} = 0. \quad (12)$$

Since $\chi_{AD}(a)$ is concave, the statement follows if we show that $\chi'_{AD}(\frac{1}{2}) > 0$. But, if $a = \frac{1}{2}$, $x = \sqrt{1 - \gamma + \gamma^2}$ and

$$\chi'_{AD}(a = 0.5) = -(1 - \gamma) \ln \frac{1 + \gamma}{1 - \gamma} + \frac{\gamma(1 - \gamma)}{x} \ln \frac{1 + x}{1 - x} > 0 \quad (13)$$

because $x > \gamma$ and the function $\frac{1}{2x} \ln \frac{1+x}{1-x} = \frac{\tanh^{-1}(x)}{x}$ is increasing. The resulting capacity is plotted in Figure 2.

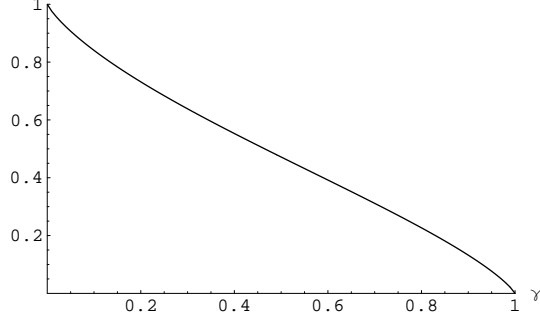


FIG. 2: The classical capacity of the qubit amplitude damping channel plotted as a function of γ .

III. CONVEX COMBINATIONS OF TWO MEMORYLESS CHANNELS

Let us now consider a convex combination of two memoryless channels. It was shown in [1] that the product-state capacity is given by Equation (4). Note that we always have

$$C_p(\Phi^{(n)}) \leq \bigwedge_{i=1}^M \left[\sup_{\{p_j, \rho_j\}} \chi_i(\{p_j, \rho_j\}) \right]. \quad (14)$$

We now consider three cases: a convex combination of two depolarizing channels, two amplitude-damping channels, and one depolarizing and one amplitude-damping channel.

A. Two depolarizing channels

In the case of a convex combination of two depolarizing qubit channels $\Delta_{\lambda_i}(\rho) = (1 - \lambda_i)\rho + \lambda_i(\frac{I}{2})$ with parameters λ_1 and λ_2 , we have

$$C(\Phi_{\lambda_1, \lambda_2}^{(n)}) = \chi^*(\lambda_1) \wedge \chi^*(\lambda_2) = \chi^*(\lambda_1 \vee \lambda_2). \quad (15)$$

Indeed, since the maximizing ensemble for both channels is the same, namely two projections onto orthogonal states, this also maximizes the minimum $\chi_1 \wedge \chi_2$. (The product-state capacity of a depolarizing qubit channel is well-known of course, and is given by $\chi^*(\Delta_\lambda) = 1 - H(\frac{\lambda}{2})$. In fact, it was proved by King [12], that this is also the classical (ultimate) capacity of the channel.)

B. Two amplitude-damping channels

A convex combination of amplitude-damping channels is similar. In that case, the maximizing ensemble does depend on the parameter γ , but as can be seen from Figure 3, for any a , $\chi_{AD}(a)$ decreases with γ , so $\chi(\gamma_1) \wedge \chi(\gamma_2) = \chi(\gamma_1 \vee \gamma_2)$ and we have again,

$$C_p(\Phi_{\gamma_1, \gamma_2}^{(n)}) = \chi^*(\gamma_1) \wedge \chi^*(\gamma_2) = \chi^*(\gamma_1 \vee \gamma_2). \quad (16)$$

In fact, for $\gamma \leq \frac{1}{2}$ this can be seen as follows. The derivative with respect to γ is given by

$$\frac{\partial \chi}{\partial \gamma} = -(1-a) \ln \frac{a + \gamma(1-a)}{(1-\gamma)(1-a)} + \frac{(2\gamma-1)(1-a)^2}{x} \ln \frac{1+x}{1-x}. \quad (17)$$

Clearly, if $\frac{a}{1-a} > 1 - 2\gamma$ both terms are negative. Otherwise, we remark that $x \geq (1-2\gamma)(1-a)$ so that it suffices if $x > y = 1 - 2\gamma - 2a(1-\gamma) > 0$. This is easily checked.

In case $\gamma > \frac{1}{2}$, we need to show that

$$f(a, \gamma) = \ln \frac{a + \gamma(1-a)}{(1-\gamma)(1-a)} - \frac{(2\gamma-1)(1-a)}{x} \ln \frac{1+x}{1-x} \geq 0.$$

Now, if $a = 0$, then $f(0, \gamma) = 0$, and the derivative is given by

$$\frac{\partial f(a, \gamma)}{\partial a} = \frac{1 - \gamma}{a + \gamma(1 - a)} + \frac{1}{1 - a} + \frac{2\gamma - 1}{x^3} \ln \frac{1 + x}{1 - x} - \frac{2(2\gamma - 1)}{x^2} \quad (18)$$

which can be shown to be positive.

C. A depolarizing channel and an amplitude-damping channel

We now investigate the product-state capacity of a convex combination of an amplitude-damping and a depolarizing channel. Let χ_1 and χ_2 denote the Holevo quantity of the amplitude-damping and depolarizing channels respectively.

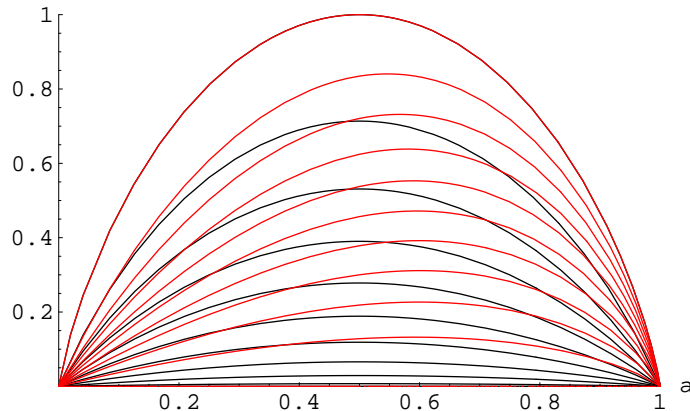


FIG. 3: The Holevo χ quantity for the amplitude damping channel and the depolarizing channel plotted as a function of a for different parameter values. The amplitude-damping channel is represented in bold.

They are plotted in Figure 3 for $0 \leq \gamma, \lambda \leq 1$. The plot above indicates that, for certain values of γ and λ the maximizer for the amplitude-damping channel lies to the right of the intersection of $\chi_1(a)$ and $\chi_2(a)$ for the depolarizing channel, whereas that for the depolarizing channel lies to the left. Indeed, keeping λ fixed, we can increase γ until the maximum of $\chi_{AD}(\gamma)$ lies above the graph of χ_{Dep} . The two graphs then intersect at a value of a intermediate between $\frac{1}{2}$ and the maximizer for χ_{AD} . This proves that the maximum of the minimum of the channels is in general not equal to the minimum of the individual channel capacities.

-
- [1] N. Datta and T.C. Dorlas, The coding theorem for a class of quantum channels with long-term memory, *Journal of Physics A, Math. Theor.* **40** (2007) 8147–8164.
 - [2] A.S. Holevo, The capacity of the quantum channel with general signal states, *IEEE Transactions on Information Theory* **44** (1998) 269–273.
 - [3] B. Schumacher and M. Westmoreland, Sending classical information via noisy quantum channels, *Phys. Rev. A* **56** (1997) 131–138.
 - [4] B. Schumacher, Sending entanglement through noisy quantum channel, *Phys. Rev. A* **54** (1996) 2614–2628.
 - [5] C. Fuchs, C. Bennett and J. Smolin, Entanglement-enhanced classical communication on a noisy quantum channel, arxiv-ph/9611006
 - [6] E.B. Davies, Information and quantum measurement, *IEEE Transactions on Information Theory* **24** (1978) 596–599.
 - [7] H.G. Eggleston, *Convexity* (Cambridge University Press, 1958).
 - [8] B. Grunbaum, *Convex Polytopes* (Interscience Publishers, 1967).
 - [9] The maximizer for this case has also been obtained in [10], but their proof is different.
 - [10] V. Gioannetti and R. Fazio, Information-capacity description of spin-chain correlations *Phys. Rev. A* **71** (2005) 032314.
 - [11] C. Fuchs, Nonorthogonal quantum states maximize classical information capacity, *Phys. Rev. Lett.* **79** (1997) 1162–1165.
 - [12] C. King, The capacity of the quantum depolarizing channel, *IEEE Transactions on Information Theory* **49** (2003) 221–229.